JOURNAL OF MULTIVARIATE ANALYSIS **59**, 13–21 (1996)
ARTICLE NO. 0051

On the Multivariate Compound Distributions*

Y. H. Wang

Concordia University, Montreal, Canada

We present two methods of constructing multivariate compound distributions and investigate the corresponding infinitely divisible and compound Poisson distributions. We then show that the multivariate compound Poisson distributions can be derived as the limiting distributions of the sums of independent random vectors.

The construction of multivariate infinitely divisible (i.d.) distribution in general and multivariate Poisson distribution in particular was first discussed in Dwass and Teicher (1957). Recently their method was used by Brown and Rinott (1988) (see also Ellis (1988)) to construct a multivariate compound Poisson distribution (MCPD).

The aim of this paper is to propose two methods of constructing multivariate compound distributions (MCD). We then explore their corresponding infinitely divisible compound distributions (MIDCD) and MCPDs. Finally, to substantiate our MCPDs, we show that they can be derived as the limiting distributions of the sums of independent random vectors.

2. THE MULTIVARIATE COMPOUND DISTRIBUTIONS

Denote by ζ a counting variable taking values in $\{0, 1, 2, \ldots\}$ and by $\mathbb{U} = \{U_i : i = 1, 2, \ldots\}$ a sequence of i.i.d. random variables with c.d.f. F. Define

$$W = U_1 + \cdots + U_\zeta. \tag{1}$$

Received August 24, 1994; revised May 1996.

AMS 1991 subject classifications: primary 60E05; secondary 60G50.

Key words and phrases: multivariate compound distribution, infinitely divisible, compound Poisson, limit theorems, sum of random vectors.

* Research partially supported by NSERC of Canada Grant OGPIN 014.

0047-259X/96 $18.00

(In (1) and in many similar definitions to come, we shall take $W=0$ if $\zeta=0$.) The characteristic function (ch.f.) G of W is

$$G(s)=\Psi(g(s)), \qquad s\in R, \tag{1.a}$$

where $g(s)=\int_R e^{isu}F(du)$ is the ch.f. of U_i and $\Psi(s)=E(s^{\zeta})$ is the probability generation function (p.g.f.) of ζ.

$G=\Psi\circ g$ is the composite of Ψ and g. Therefore it is a ch.f. of a *compound distribution with compounding distribution F*. If Ψ is degenerate at 1, then $G=g$. In such a case we shall not consider G a compound distribution. Therefore we shall assume all the counting variables in this paper non-degenerate.

The multivariate i.d. distribution (MIDD) in the literature today is the one constructed by Dwass and Teicher (1957) as follows:

Let $M=\{1, 2, ..., m\}$, where $m\geqslant 2$, $\mathscr{A}=\{A: \phi\neq A\subseteq M\}$. A random vector $W=(W_1, ..., W_m)$ is said to have the MIDD if

$$W_i=\sum_{A:\, i\in A} Z_A, \qquad i=1, ..., m, \tag{2}$$

where $\{Z_A: A\in\mathscr{A}\}$ are 2^m-1 independent i.d. random variables. Its ch.f. is

$$G(s)=\prod_{A\in\mathscr{A}} g_A\left(\sum_{i\in A} s_i\right), \qquad s\in R^m, \tag{2.a}$$

where g_A denotes the ch.f. of Z_A.

If $\{Z_A: A\in A\}$ is an independent Poisson sequence, then the distribution of W defined by (2) is the multivariate Poisson distribution.

We shall use Dwass and Teicher's approach to construct two MCDs. The first one involves one single compounding distribution and one single counting variable while the second one involves several compounding distributions and several counting variables.

Definition 1. A random vector W is said to have the *multivariate compound distribution of type I* (MCD-I) with compounding distribution F if

$$W=U_1+\cdots+U_{\zeta}, \tag{3}$$

where ζ is a counting variable and independent of the i.i.d. sequence $\mathbb{U}=\{U_i: i=1, 2, ...\}$ with common c.d.f. F. The ch.f. G of W is

$$G(s)=\Psi[g(s)], \qquad s=(s_1, ..., s_m)\in R^m, \tag{3.a}$$

where Ψ is the p.g.f. of ζ and $g(s)=\int_{R^m} e^{isu}F(du)$.

DEFINITION 2. A random vector $W=(W_1, ..., W_m)$ is said to have the *multivariate compound distribution of type II* (MCD-II) with compounding distributions $\{F_A: A\in\mathscr{A}\}$ if

$$W_i = \sum_{A:\, i\in A} [U_{A1} + \cdots + U_{A,\,\zeta(A)}], \qquad i=1, ..., m, \tag{4}$$

where $\{\zeta(A): A\in\mathscr{A}\}$ are 2^m-1 independent counting variables and are independent of the double sequence of independent random variables $\mathbb{U}=\{U_{Ai}: i=1, 2, ...; A\in\mathscr{A}\}$ in which, for each $A\in\mathscr{A}$, $\{U_{Ai}\}$ is a sequence of i.i.d. random variables with common c.d.f. F_A. The ch.f. G of W is

$$G(s) = \prod_{A\in\mathscr{A}} \Psi_A\left[g_A\left(\sum_{i\in A} s_i\right)\right], \qquad s=(s_1, ..., s_m)\in R^m, \tag{4.a}$$

where $g_A(s)=\int_R e^{isu}F_A(du)$ and Ψ_A is the p.g.f. of $\zeta(A)$.

3. THE MULTIVARIATE I.D. COMPOUND DISTRIBUTIONS

The ch.f. G in (3.a) is i.d. if Ψ is. Therefore if Ψ is i.d. the MCDs become MIDCDs. We shall denote the corresponding i.d. versions of the two types of MCDs by MIDCD-I and MIDCD-II, respectively.

The following two properties of the MIDCDs can be easily verified:

(a) Inheritance property: The marginal distributions of any proper subsets of size k, $1\leqslant k<m$, of both types are of the same type. In particular, the one-dimensional marginal distribution is the univariate i.d. compound distribution (1.a). In this case, they coincide.

(b) Additivity property: If W and Y are independent random vectors of the same dimension having the MIDCD-I (MIDCD-II) with the same compounding distribution(s) F ($\{F_A: A\in\mathscr{A}\}$), then the sum $W+Y$ has also the MIDCD-I (MIDCD-II).

Aside from the above two properties, those two types of i.d. distributions seem to share no further properties. Neither is a special case of the other. The MIDD can be reduced from the MIDCD-II by taking the compounding distributions $\{F_A: A\in\mathscr{A}\}$ to be degenerate at 1. Therefore the MIDD is a special case of the MIDCD-II, but not of the MIDCD-I, which is not unexpected, since the construction of the MIDCD-II literally follows that of the MIDD.

4. THE MULTIVARIATE COMPOUND POISSON DISTRIBUTION

In the next two sections, we study the multivariate compound Poisson distribution. The compound Poisson distribution is an important sub-class of the i.d. distributions, because, as stated in Brown and Rinott (1988), "The class of limits in distribution of sequences of infinitely divisible compound Poisson distributions and the class of infinitely divisible distributions are identical."

Following (1.a), we say that a distribution is the compound Poisson distribution with parameter $\lambda(t)$ and compounding distribution F if its ch.f. is

$$G(s, t) = \exp\{\lambda(t)[g(s) - 1]\} \quad \text{for} \quad s \in R \quad \text{and} \quad \lambda(t) \geqslant 0, \tag{5}$$

where $g(s) = \int_R e^{isu} F(du)$.

Definition 3. Let $\mathbb{T} = \{t : t \geqslant 0\}$. For $t \in \mathbb{T}$, a random vector W_t is said to have the *multivariate compound Poisson distribution of type I* (MCPD-I) with parameters $\{\lambda(t) : t \in \mathbb{T}\}$ and compounding distribution F if

$$W_t = U_1 + \cdots + U_{\zeta(t)}, \tag{6}$$

where $\zeta(t)$ is Poisson with parameter $\lambda(t)$ and is independent of the sequence of i.i.d. random variables $\mathbb{U} = \{U_i\}$ with common c.d.f. F. The ch.f. of W_t is

$$h(s, t) = \exp\{\lambda(t)[g(s) - 1]\}, \qquad s \in R^m, \quad t \in \mathbb{T}, \tag{6.a}$$

where $g(s) = \int_{R^m} e^{is'x} F(dx)$.

Let $\mathbb{T}(A) = \{t(A) : t(A) \geqslant 0\}$, $A \in \mathscr{A}$, and $\mathbb{U} = \{U_{Ai}\}$ be a double array of independent random variables and independent of the Poisson sequence $\{\zeta(t(A)) : t(A) \in \mathbb{T}(A)\}$ with parameters $\{\lambda(t(A)) : t(A) \in \mathbb{T}(A)\}$ such that U_{Ai} are i.i.d. for $A \in \mathscr{A}$ with common c.d.f. F_A. Define 2^{m-1} random variables

$$Z_A = U_{A1} + \cdots + U_{A, \zeta(T(A))} \qquad \text{for} \quad A \in \mathscr{A}.$$

Definition 4. A random vector $W = (W_1, ..., W_m)$ is said to have the *multivariate Compound Poisson distribution of type II* (MCPD-II) with parameters $\{\lambda(t(A)) : t(A) \in \mathbb{T}(A)\}$ and compounding distributions F_A, $A \in \mathscr{A}$, if

$$W_i = \sum_{A:\, i \in A} Z_A, \qquad i = 1, ..., m. \tag{7}$$

The ch.f. of W is thus

$$h(s,t)=\prod_{A\in A}\exp\left\{t(A)\left[f_A\left(\sum_{i\in A}s_i\right)-1\right]\right\},\qquad s\in R^m,\quad t\geqslant 0 \quad (7.a)$$

where $f_A(s)=\int_R e^{isx}F_A(dx)$.

The two types of MCPDs are not necessarily i.d. They are i.d. if and only if $\lambda(t)=\lambda t$ for the MCPD-I and $\lambda(t(A))=\lambda(A)\,t$ for the MCPD-II.

As in the MIDCDs case, the MCPDs also possess the inheritance and the additivity properties. If the compounding distributions $\{F_A\colon A\in\mathscr{A}\}$ are all degenerate at 1, the MCPD-II reduces to the multivariate Poisson distribution (2). If $m=1$, both distributions reduce to the same univariate compound Poisson distribution (5).

The multivariate compound Poisson distribution constructed by Brown and Rinott (1988) engages one single compounding distribution while the MCPD-II involves 2^m-1 compounding distributions.

5. MCPD AS LIMITING DISTRIBUTIONS

To substantiate our constructions of the MCPD, we derive the two types of MCPDs as the limiting distributions of sums independent random vectors. For brevity, we shall write $\lambda(A)$ for $\lambda(t(A))$ and λ for $\lambda(t)$.

The proof of the following lemma can be found in Wang (1989a).

LEMMA 5. *Let θ be a complex number such that $|\theta|<1$, then*

$$-\ln(1-\theta)=\theta+\theta^2K(\theta),$$

where $|K(\theta)|<1$ if $|\theta|<1/2$.

Let F be a c.d.f. in R^m and $\{(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)\}$ be $n\geqslant 2$ $(m+1)$-dimensional independent random vectors such that the one-dimensional marginal distributions of $\{Y_j\}$ are Bernoulli with $P(Y_j=1)=p_j$ and the conditional distributions of the random vectors X_j given Y_j are

$$P(X_j=0\mid Y_j=0)=1\qquad\text{and}\qquad P(X_j\leqslant x\mid Y_j=1)=F(x)\qquad (8)$$

for all $j=1, ..., n$ and $x\in R^m$.

The c.d.f. G_j of X_j in (8) can be written in a mixture form as $G_j=p_jF+(1-p_j)\,H$, where H denotes a degenerate c.d.f. at 0 in R^m.

A motivation for the above model can be put as follows: Consider automobile insurance as an example. Divide a time interval into n mutually exclusive and exhaustive *small* subintervals $I_1, ..., I_n$, such that in each interval the probability of having two or more accidents is negligible, while that of having exactly one accident, even though not negligible, is very small. We let Y_i be the indicator variable of I_i, and X_i be the relevant information about the accident occurred in the interval I_i, such as the driver's age, the number of occupants in the car, the amount of compensation claimed, etc. If no accidents occur in an interval, nothing is registered. If one accident occurs, then the distribution of the relevant information is $F(x)$.

Theorem 6. *If*

$$\sum_{j=1}^{n} p_j \to \lambda > 0 \qquad \sum_{j=1}^{n} p_j^2 \to 0 \qquad \textit{as} \quad n \to \infty$$

then the limiting distribution of $S_n = X_1 + \cdots + X_n$ *is the* MCPD-I *with parameter* λ *and compounding distribution* F.

Proof. Let $h_{nj}(s) = E[\exp(is'X_j)]$. Then conditioning on Y_i and by (8)

$$h_{nj}(s) = 1 - p_j[1 - f(s)], \qquad s \in R^m,$$

where $f(s) = \int_{R^m} e^{is'y} F(dy)$.

Since X_j's are independent, the logarithm of the ch.f. h_n of S_n, for sufficiently large n, is

$$\begin{aligned} -\ln[h_n(s)] &= -\sum_{j=1}^{n} \ln[1 - p_j\{1 - f(s)\}] \\ &= [1 - f(s)] \sum_{j=1}^{n} p_j + [1 - f(s)]^2 \sum_{j=1}^{n} p_j^2 K[p_j\{1 - f(s)\}]. \end{aligned} \tag{9}$$

The last equality in (9) follows from Lemma 5 by choosing n large enough so that $\max_{1 \leqslant j \leqslant n} p_j < 1/4$. The absolute value of the last term in (9) is bounded above by $[1 - f(s)]^2 \sum_{j=1}^{n} p_j^2$ which tends to 0 as $n \to \infty$ for all $s \in R^m$.

This proves

$$\lim_{n \to \infty} h_n(s) = \exp\{\lambda[f(s) - 1]\}, \qquad s \in R^m.$$

For each $A \in \mathscr{A}$, let F_A be a c.d.f. on R and H be a c.d.f. degenerate at $0 \in R$. Let $\{Y_{Ak}: k = 1, ..., n;\ A \in \mathscr{A}\}, n \geqslant 2$, be a sequence of independent random variables with distribution functions

$$\{G_{Ak}: G_{Ak} = p_{nk}(A) F_A + (1 - p_{nk}(A)) H, k = 1, ..., n;\ A \in \mathscr{A}\}. \tag{10}$$

For each n and k, define n m-dimensional independent random vectors $X_k = (X_{1k}, ..., X_{mk})$, for $k = 1, ..., n$, by

$$X_{jk} = \sum_{A:\, j \in A} Y_{Ak}, \qquad \text{for} \quad j = 1, ..., m \qquad \text{and} \qquad k = 1, ..., n. \tag{11}$$

Denote

$$\lambda_n(A) = \sum_{k=1}^{n} p_{nk}(A) \qquad \text{and} \qquad M_n = \max_{A \in \mathscr{A}} \sum_{k=1}^{n} p_{nk}^2(A).$$

THEOREM 7. *If $\lambda_n(A) \to \lambda(A) > 0$ for all $A \in \mathscr{A}$ and $M_n \to 0$ as $n \to \infty$, then the limiting distribution of $S_n = X_1 + \ldots + X_n$ is the* MCPD-II *with parameters $\{\lambda(A): A \in \mathscr{A}\}$ and compounding distributions $\{F_A: A \in \mathscr{A}\}$.*

Proof. Let

$$h_{nk}(s) = E\left[\exp\left(i \sum_{j=1}^{m} s_j X_{jk}\right)\right], \qquad s \in R^m,$$

be the ch.f. of X_k, then by (11)

$$h_{nk}(s) = \prod_{A \in \mathscr{A}} \left\{1 - p_{nk}(A)\left[1 - f_A\left(\sum_{j \in A} s_j\right)\right]\right\},$$

where $f_A(s) = \int_R e^{isx} F_A(dx)$.

Hence, the ch.f. of S_n is

$$h_n(s) = \prod_{k=1}^{n} \left\{\prod_{A \in A} \left\{1 - p_{nk}(A)\left[1 - f_A\left(\sum_{j \in A} s_j\right)\right]\right\}\right\}. \tag{12}$$

Since, for all $A \in \mathscr{A}$ and k,

$$p_{nk}(A) \leqslant M_n^{1/2} \to 0 \qquad \text{as} \quad n \to \infty,$$

we shall, without loss of generality, assume $M_n < 1/16$ for all n so that $p_{nk}(A) < 1/4$ for all $A \in \mathscr{A}$ and $k = 1, ..., n$.

Taking the logarithm on both sides of (12) and using Lemma 5 we get

$$-\ln[h_n(s)] = \sum_{A \in \mathscr{A}} \left[1 - f_A\left(\sum_{j \in A} s_j\right)\right] \lambda_n(A) + R_n,$$

where the remainder term R_n is

$$R_n = \sum_{A \in \mathscr{A}} \left[1 - f_A\left(\sum_{j \in A} s_j\right)\right]^2 \sum_{k=1}^{n} p_{nk}^2(A)\, K[p_{nk}(A)\left[1 - f_A\left(\sum_{j \in A} s_j\right)\right].$$

Therefore, $|R_n| \leqslant 4M_n \to 0$ as $n \to \infty$, and we conclude that for all $s \in R^m$,

$$\lim_{n \to \infty} \ln[h_n(s)] = \sum_{A \in \mathscr{A}} \lambda(A) \left[f_A\left(\sum_{j \in A} s_j \right) - 1 \right].$$

The classical Poisson convergence theorem deals with the discrete case. To date most of the results on this topic are also confined to this case. The main reason could be that the discrete case offers some inside view of the Poisson convergence theorem which seems lacking in the continuous case, even for the compound Poisson theorems. In the following we give two theorems for the discrete multivariate compound Poisson distribution. They can be proved along the lines of the proofs for Theorems 6 and 7 and shall be skipped.

Let $\{X_k : k = 1, ..., n\}$ be n independent discrete random vectors taking values in a common set $D \subseteq \{0, \pm 1, \pm 2, ...\}^m$ such that $0 \in D$. Denote $p_k(x) = P(X_k = x)$ for $x \in D$; $\lambda_n(x) = \sum_{k=1}^n p_k(x)$ for $x \in D \setminus \{0\}$; $\lambda_n = \sum_{x \in D} \setminus \{0\} \lambda_n(x)$; and $\theta_k = 1 - p_k(0) > 0$ for all k and $M_n = \sum_{k=1}^n \theta_k^2$.

THEOREM 8. *If* $\lambda_n \to \lambda > 0, \lambda_n(x) \to \lambda(x) \geqslant 0$ *for all* $x \in D \setminus \{0\}$ *and* $M_n \to 0$, *as* $n \to \infty$, *then the limiting distribution of* $S_n = X_1 + \cdots + X_n$ *is the* MCPD-I *with parameter* λ *and compounding distributions* $\{r(x) : x \in D \setminus \{0\}\}$, *where* $r(x) = \lambda(x)/\lambda$ *for* $x \in D \setminus \{0\}$.

Let $\{Y_{Ak} : k = 1, ..., n; A \in \mathscr{A}\}, n \geqslant 2$, be a double sequence of independent discrete random variables taking values in $D \subseteq \{0, \pm 1, \pm 2, ...\}$ such that $0 \in D$ and with probability mass function $p_{Ak}(x) = P(Y_{Ak} = x)$. Let

(a) $\quad \lambda_{An}(x) = \sum_{k=1}^n p_{Ak}(x) \qquad$ for $\quad x \in D \setminus \{0\}$;

(b) $\quad \Lambda_{An} = \sum_{x \in D \setminus \{0\}} \lambda_{An}(x)$;

(c) $\quad \theta_{Ak} = 1 - p_{Ak}(0) > 0 \qquad$ for all k;

(d) $\quad M_n = \max_A \sum_{k=1}^n \theta_{Ak}^2$.

Define n m-dimensional independent random vectors $X_k = (X_{1k}, ..., X_{mk})$ by

$$X_{jk} = \sum_{A : j \in A} Y_{Ak}, \qquad j = 1, ..., m; \quad k = 1, ..., n.$$

THEOREM 9. *If, for each* $A \in \mathscr{A}$,

$$\Lambda_{An} \to \Lambda_A > 0, \lambda_{An}(x) \to \lambda_A(x) \geqslant 0, \qquad \text{and} \qquad M_n \to 0, \qquad \text{as} \quad n \to \infty,$$

then the limiting distribution of $S_n = X_1 + \ldots + X_n$ *is the* MCPD-II *with parameters* $\{\lambda_A : A \in \mathscr{A}\}$ *and compounding distributions* $\{r_A(x);\ x \in D \backslash \{0\};\ A \in \mathscr{A}\}$, *where* $r_A(x) = \lambda_A(x)/\Lambda_A$, *for* $x \in D\backslash\{0\}$.

A final remark: In Theorem 9, if we take $D = \{0, 1\}$, then the compounding distribution r_A is degenerate at $x = 1$, for all $A \in \mathscr{A}$. The limit is the multivariate Poisson distribution in the sense of Dwass and Teicher. (See Wang (1974) for a different derivation of this theorem.) If, in addition, $m = 1$, and Y_k's are i.i.d., then it reduces to the classical Poisson convergence theorem.

REFERENCES

Brown, L. D., and Rinott, Y. (1988). Inequalities for multivariate infinitely divisible processes. *Ann. Probab.* **16**, 642–657.

Dwass, M., and Teicher, H. (1957). On infinitely divisible random vectors. *Ann. Math. Statist.* **28**, 461–470.

Ellis, R. S. (1988). Inequalities for multivariate compound Poisson distributions. *Ann. Probab.* **16**, 658–661.

Wang, Y. H. (1974). Characterization of certain multivariate distributions. *Proc. Cambridge Philos. Soc.* **75**, 219–234.

Wang, Y. H. (1989a). From Poisson to compound Poisson approximations. *Math. Sci.* **14**, 69–74.

Wang, Y. H. (1989b). A multivariate extension of Poisson's theorem. *Canad. J. Statist.* **17**, 241–245.